JOURNAL OF COMPLEXITY 2, 239–254 (1986)

Optimal Complexity Recovery of Band- and Energy-Limited Signals

MAREK A. KOWALSKI

Institute of Informatics, University of Warsaw, 00-901 Warsaw, PKin 850, Poland

Received December 31, 1985

This paper deals with the recovery of band- and energy-limited signals from a finite set of their samples taken in a given finite interval. Let $m(\epsilon)$ be the minimal number of samples required to get an ϵ-accurate approximation of any such signal. We prove that

$$\lim_{\epsilon\to 0^+} \frac{m(\epsilon)\log\log 1/\epsilon}{\log 1/\epsilon} = 1,$$

and, for sufficiently small $\epsilon > 0$, Lagrangian interpolation with $m(\epsilon)(1 + o(1))$ arbitrary nodes yields an ϵ-approximation with almost minimal cost.

1. INTRODUCTION

This paper deals with the recovery of signals $\check{X}$ of bandwidth $[-\Omega_0, \Omega_0]$,

$$\check{X}(t) \equiv \int_{-\Omega_0}^{\Omega_0} X(\Omega)\exp(i\Omega t)d\Omega \qquad (X \in L_2(-\Omega_0, \Omega_0),\ i = \sqrt{-1}),$$

from a set of their samples taken in a given finite interval $I = [a - \tau, a + \tau]$. That is, we want to approximate $\check{X}$ with our sole knowledge of $\check{X}$ being information of the form

$$N^{a}(\check{X}) = [\check{X}(\vartheta_1), \check{X}(\vartheta_2), \ldots, \check{X}(\vartheta_n)]^{T},$$

where adaptive choice of the sampling points $\vartheta_k \in I$ is allowed, i.e.,

$$\vartheta_k = \vartheta_k(\check{X}(\vartheta_1), \ldots, \check{X}(\vartheta_{k-1})), \qquad k = 2, 3, \ldots, n.$$

0885-064X/86 $3.00

For a practical approximation scheme, n must be finite. Thus $N^a(\check{X})$ never determines $\check{X}$. In fact matters are even worse, since $N^a(\check{X})$ does not determine $\check{X}$ with a bounded accuracy (measured by any norm). To obtain a meaningful problem, we take a uniform bound on the energy of the signals to be considered. For simplicity we choose the bound 2π. By the Parseval theorem we have

$$\int_{-\infty}^{+\infty} |\check{X}(t)|^2\,dt \le 2\pi \qquad \text{iff } \|X\|^2 \equiv \int_{-\Omega_0}^{\Omega_0} |X(\Omega)|^2 d\Omega \le 1.$$

Hence the resulting class of signals is

$$J_0 = \{\check{X} : X \in B(\Omega_0)\},$$

where

$$B(\Omega_0) = \{Y \in L_2(-\Omega_0, \Omega_0) : \|Y\| \le 1\}.$$

We study the recovery of $\check{X} \in J_0$ from its samples in the worst-case setting (see Traub and Woźniakowski, 1980).

It is known that the adaptive information $N^a(\check{X})$ is not more powerful than the corresponding nonadaptive information

$$N\check{X} = [\check{X}(t_1), \check{X}(t_2), \dots, \check{X}(t_n)]^T,$$

where

$$t_1 = \vartheta_1, \qquad t_k = \vartheta_k(0, 0, \dots, 0), \qquad k = 2, 3, \dots, n$$

(see Traub and Woźniakowski, 1980). Thus we assume that the sampling points are simultaneously (nonadaptively) chosen.

Given $\epsilon > 0$, let $m(\epsilon)$ denote the minimal number of samples required to find a set of functions $\{a_{\epsilon,\check{X}}\}_{\check{X}\in J_0}$ satisfying

$$\sup\{|\check{X}(t) - a_{\epsilon,\check{X}}(t)| : t \in I, \check{X} \in J_0\} \le \epsilon.$$

We call $a_{\epsilon,\check{X}}(t)$ an *ϵ-approximation* to $\check{X}(t)$. Assume that the cost of the arithmetic operations $(+, -, \times, /)$ and the cost of the signal evaluation are taken as unity and c, respectively. Let $\text{comp}(\epsilon)$ be the minimal computing cost (complexity) of $a_{\epsilon,\check{X}}(t)$.

The main result of this paper is the following theorem.

THEOREM 1.

$$\text{(i)} \quad \lim_{\epsilon\to 0^+} \frac{m(\epsilon)\log\log 1/\epsilon}{\log 1/\epsilon} = 1.$$

(ii) $\operatorname{comp}(\epsilon) = \Theta\left(\dfrac{\log 1/\epsilon}{\log\log 1/\epsilon}\right)$ as $\epsilon \to 0^+$.

(iii) *For sufficiently small* $\epsilon > 0$, *Lagrangian interpolation with* $m(\epsilon)(1 + o(1))$ *arbitrary distinct nodes from* I *yields an* ϵ*-approximation with almost minimal cost.*

The Θ-notation used in (ii) can be thought as a "two-sided" O-notation, i.e., $p = \Theta(q)$ iff $p = O(q)$ and $q = O(p)$.

2. Auxiliary Lemmas and Remarks

Let $t_1, t_2, \ldots, t_n$ be arbitrary distinct points from the interval I. Define

$$u_k(\Omega) = \exp(-it_k\Omega), \qquad k = 1, 2, \ldots, n,$$

and denote by $G = G(u_1, u_2, \ldots, u_n)$ the Gram matrix $(\langle u_l, u_k\rangle)_{k,l=1}^n$, where $\langle\ ,\ \rangle$ is the inner product in $L_2(-\Omega_0, \Omega_0)$.

Suppose that for any $\check{X} \in J_0$ and any $t \in I$ we wish to recover $\check{X}(t)$ as good as possible from

$$N\check{X} = [\check{X}(t_1), \check{X}(t_2), \ldots, \check{X}(t_n)]^{\mathrm{T}} = [\langle X, u_1\rangle, \langle X, u_2\rangle, \ldots, \langle X, u_n\rangle]^{\mathrm{T}}. \quad (1)$$

That is, we are looking for a mapping (algorithm)

$$\varphi\colon N(J_0) \to \mathbf{V} \equiv \{f : f\colon I \to \mathbf{C}\},$$

which minimizes the worst-case error

$$e(\varphi) = \sup\{e(\varphi; t) : t \in I\},$$
$$e(\varphi; t) = \sup\{|\check{X}(t) - \varphi(N\check{X})(t)| : \check{X} \in J_0\}.$$

Define

$$r(t) \equiv r(t; t_1, t_2, \ldots, t_n) = \sup\{|\check{X}(t)| : \check{X} \in J_0, N\check{X} = 0\} \quad (2)$$

and

$$R \equiv R(t_1, t_2, \ldots, t_n) = \sup\{r(t) : t \in I\}.$$

As a consequence of Micchelli and Rivlin (1977, Theorem 3, Ex. 1.1) we have

$$r(t) = \inf_{\varphi} e(\varphi; t) \qquad \text{and} \qquad R = \inf_{\varphi} e(\varphi),$$

where the infima are taken over all algorithms φ: $N(J_0) \to \mathbf{V}$ and achieved by

$$\varphi^*: \varphi^*(N\check{X}) = \sum_{k=1}^{n} (G^{-1}N\check{X})_k \check{u}_k.$$

Alternately, we may write

$$\varphi^*(N\check{X}) = \sum_{k=1}^{n} a_k \mathrm{sa}(\Omega_0(t_k - \cdot)),$$

where

$$\mathrm{sa}(x) = \begin{cases} \sin(x)/x & \text{if } x \neq 0, \\ 1 & \text{if } x = 0 \end{cases}$$

and coefficients a_k are determined by the linear system

$$\sum_{k=1}^{n} a_k \mathrm{sa}(\Omega_0(t_j - t_k)) = \check{X}(t_j), \qquad j = 1, 2, \ldots, n.$$

We remark that if $t_k = (k_0 + k)\pi/\Omega_0$, $k = 1, 2, \ldots, n$, then $\varphi^*(N\check{X})(t)$ is a partial sum of the Whittaker cardinal series

$$\check{X}(t) \equiv \sum_{k=-\infty}^{\infty} \check{X}(k\pi/\Omega_0)\mathrm{sa}(k\pi - t\Omega_0)$$

(see Butzer, 1983). Note that in this case the error of φ^* is large, since

$$e(\varphi^*) \geq r(t_n - \pi/2\Omega_0) \geq (2\Omega_0)^{-1/2}\, \widecheck{\exp(-it_n \cdot)}(t_n - \pi/2\Omega_0) \\ = 2(2\Omega_0)^{1/2}/\pi$$

(see Landau, 1985, for related results).

To prove Theorem 1 we need a few lemmas.

Fix $t_0 \in I \backslash \{t_1, t_2, \ldots, t_n\}$ and define

$$u_0(\Omega) \equiv \exp(-it_0\Omega),$$
$$r \equiv r(t_0; t_1, t_2, \ldots, t_n)$$

and

$$M \equiv (2\Omega_0)^{-1}G(u_0, u_1, \ldots, u_n) = (\mathrm{sa}(\Omega_0(t_l - t_k)))_{l,k=0}^{n}.$$

Let $(M^{-1})_{oo}$ denote the first diagonal coefficient of the matrix M^{-1}.

LEMMA 1. $r^2 = 2\Omega_0/(M^{-1})_{oo}$.

Proof. By (1), (2), and $\check{X}(t_0) = \langle X, u_0\rangle$ we get

$$r = \sup\{|\langle X, u_0\rangle| : X \in B(\Omega_0) \cap \operatorname{span}(u_1, u_2, \ldots, u_n)^{\perp}\} = \|Pu_0\|,$$

where Pu_0 is the orthogonal projection of u_0 on $\operatorname{span}(u_1, u_2, \ldots, u_n)^{\perp}$. Thus, r measures how accurately u_0 can be approximated by functions from $\operatorname{span}(u_1, u_2, \ldots, u_n)$, i.e.,

$$r = \inf\left\{\left\| u_0 - \sum_{k=1}^{n} c_k u_k \right\| : c_k \in \mathbf{C}\right\}.$$

In its dual version, this is the problem of finding the norm $\|\mathbf{L}\|$ of the unique linear functional $\mathbf{L}$ on $G_n = \operatorname{span}(u_0, u_1, \ldots, u_n)$ satisfying

$$\mathbf{L}(u_0) = 1 \qquad \text{and} \qquad \mathbf{L}(u_k) = 0 \qquad \text{for} \qquad k = 1, 2, \ldots, n. \tag{3}$$

More precisely, we have

$$r = 1/\|\mathbf{L}\| \tag{4}$$

(see Kowalski and Sawoń, 1983, or Krein and Nudelman, 1973). By the Riesz representation theorem there exists a unique function $g_{\mathbf{L}} \in G_n$ such that

$$\mathbf{L}(f) = \langle g_{\mathbf{L}}, f\rangle \qquad \text{and} \qquad \|\mathbf{L}\| = \|g_{\mathbf{L}}\|.$$

Thus, (3) can be rewritten in the form

$$\langle g_{\mathbf{L}}, u_0\rangle = 1 \qquad \text{and} \qquad \langle g_{\mathbf{L}}, u_k\rangle = 0 \qquad \text{for} \qquad k = 1, 2, \ldots, n. \tag{5}$$

This means that the coefficients a_k such that $g_{\mathbf{L}} = \sum_{k=0}^{n} a_k u_k$ satisfy

$$[a_0, a_1, \ldots, a_n]G(u_0, u_1, \ldots, u_n) = [1, 0, \ldots, 0].$$

This and (5) yield

$$\begin{aligned}\|\mathbf{L}\|^2 = \|g_{\mathbf{L}}\|^2 &= \left(\sum_{k=0}^{n} a_k u_k, g_{\mathbf{L}}\right) = a_0 = (G(u_0, u_1, \ldots, u_n)^{-1})_{oo} \\ &= (M^{-1})_{oo}/2\Omega_0.\end{aligned}$$

Applying (4), we complete the proof. ∎

To simplify notation, in what follows we shall write $\Sigma'^{u}_{k=l}$ instead of $\Sigma^{u}_{k=l,k\neq 0}$, $u > 0 > l$.

LEMMA 2.

$$\frac{\mathrm{sa}(x-y)}{\mathrm{sa}(x)\mathrm{sa}(y)} = 1 + \sum_{k=-\infty}^{\infty}{}' \frac{xy}{(x-k\pi)(y-k\pi)}, \qquad \forall\, x, y \in \mathbf{C}.$$

Proof. Let

$$F(x, y) = 1 + \sum_{k=-\infty}^{\infty}{}' \frac{xy}{(x-k\pi)(y-k\pi)}.$$

Assume without loss of generality that $x \neq y$ and $x, y \notin \{k\pi : k = \pm 1, \pm 2, \ldots\}$. Then

$$\begin{aligned} F(x, y) &= 1 + xy \lim_{m\to\infty} \sum_{k=-m}^{m}{}' \frac{1}{(x-k\pi)(y-k\pi)} \\ &= 1 + \frac{xy}{y-x} \lim_{m\to\infty} \sum_{k=-m}^{m}{}' \left(\frac{1}{x-k\pi} - \frac{1}{y-k\pi}\right) \\ &= 1 + \frac{xy}{y-x} \lim_{m\to\infty} \sum_{k=1}^{m} \left(\frac{1}{x-k\pi} + \frac{1}{x+k\pi} - \frac{1}{y-k\pi} - \frac{1}{y+k\pi}\right) \\ &= 1 + \frac{xy}{y-x}\left(2x \sum_{k=1}^{\infty} \frac{1}{x^2-k^2\pi^2} - 2y \sum_{k=1}^{\infty} \frac{1}{y^2-k^2\pi^2}\right). \end{aligned}$$

Using the Mittag–Leffler formula (see Saks and Zygmund, 1971, p. 310),

$$2z \sum_{k=1}^{\infty} \frac{1}{z^2-k^2\pi^2} = \mathrm{ctg}(z) - 1/z,$$

we get

$$F(x, y) = \frac{xy}{y-x}(\mathrm{ctg}\, x - \mathrm{ctg}\, y) = \frac{\mathrm{sa}(x-y)}{\mathrm{sa}(x)\mathrm{sa}(y)},$$

as claimed. ∎

Now define

$$s_k \equiv \frac{t_k - t_0}{\tau}, \qquad k = 0, 1, 2, \ldots, n, \ \kappa \equiv \pi/\Omega_0\tau.$$

LEMMA 3.

$$(M^{-1})_{oo} = 1 + \inf\left\{ \sum_{k=-\infty}^{\infty}{}' a_k^2 : \sum_{k=-\infty}^{\infty}{}' a_k \frac{1}{s_l - k\kappa} = 1/s_l,\ l = 1, \ldots, n \right\}.$$

Proof. Note that $(M^{-1})_{oo} = (\mathcal{M}^{-1})_{oo}$, where

$$\mathcal{M} \equiv \left(\frac{\mathrm{sa}(\Omega_0 \tau(s_l - s_k))}{\mathrm{sa}(\Omega_0 \tau s_l)\mathrm{sa}(\Omega_0 \tau s_k)} \right)_{l,k=0}^{n}.$$

By Lemma 2 we get

$$\mathcal{M} = \left(1 + \sum_{j=-\infty}^{\infty} \frac{s_l s_k}{(s_l - j\kappa)(s_k - j\kappa)} \right)_{l,k=0}^{n} = \mathbf{E}\mathbf{E}^{\mathrm{T}},$$

where

$$\mathbf{E} = \begin{bmatrix} \cdots & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & \cdots \\ \cdots & \dfrac{s_1}{s_1 + j\kappa} & \cdots & \dfrac{s_1}{s_1 + \kappa} & 1 & \dfrac{s_1}{s_1 - \kappa} & \cdots & \dfrac{s_1}{s_1 - j\kappa} & \cdots \\ & & & & \vdots & & & & \\ \cdots & \dfrac{s_n}{s_n + j\kappa} & \cdots & \dfrac{s_n}{s_n + \kappa} & 1 & \dfrac{s_n}{s_n - \kappa} & \cdots & \dfrac{s_n}{s_n - j\kappa} & \cdots \end{bmatrix}.$$

Thus, $(M^{-1})_{oo} = e_1^{\mathrm{T}}(\mathbf{E}\mathbf{E}^{\mathrm{T}})^{-1}e_1$ with $e_1 = [1, 0, \ldots, 0]^{\mathrm{T}} \in \mathbf{R}^n$. Note that

$$\begin{aligned} e_1^{\mathrm{T}}(\mathbf{E}\mathbf{E}^{\mathrm{T}})^{-1}e_1 &= \sup\{e_1^{\mathrm{T}}y - \tfrac{1}{4}y^{\mathrm{T}}\mathbf{E}\mathbf{E}^{\mathrm{T}}y : y \in \mathbf{R}^n\} \\ &= \sup\{a^{\mathrm{T}}a + (\mathbf{E}a - e_1)^{\mathrm{T}}y : 2a + \mathbf{E}^{\mathrm{T}}y = 0, y \in \mathbf{R}^n, \\ &\qquad a = [\ldots, a_{-1}, a_0, a_1, \ldots]^{\mathrm{T}} \in l_2\} \\ &= \sup\{\inf\{a^{\mathrm{T}}a + (\mathbf{E}a - e_1)^{\mathrm{T}}y : a \in l_2\} : y \in \mathbf{R}^n\} \\ &= \inf\{a^{\mathrm{T}}a : \mathbf{E}a = e_1, a \in l_2\}. \end{aligned}$$

(Consult Hestens (1966) for the third equality and Golshtein (1971) for the last dual equality.) Since $\mathbf{E}a = e_1$ implies that $a_0 = 1$ we finally get

$$(M^{-1})_{oo} = 1 + \inf\left\{ \sum_{k=-\infty}^{\infty}{}' a_k^2 : \sum_{k=-\infty}^{\infty}{}' a_k \frac{1}{s_l - k\kappa} = 1/s_l, \right.$$
$$\left. l = 1, 2, \ldots, n \right\},$$

as required. ■

We are now in a position to estimate $R(t_1, t_2, \ldots, t_n)$.

LEMMA 4.

$$R(t_1, t_2, \ldots, t_n) > \frac{(\Omega_0/n^*)^{1/2}}{(4 + \kappa n^*)^{n^*}},$$

where

$$n^* = \begin{cases} n & \text{if } n \text{ is even,} \\ n+1 & \text{if } n \text{ is odd.} \end{cases}$$

Proof. Suppose first that $n = 2m$. It is well known that

$$\sup\{|w(x)| : x \in [-1, +1]\} \geq 2$$

for any real polynomial w of the form $w(x) = \Pi_{k=1}^{n} 2(\zeta_k - x)$; see Meinardus (1964). Thus, there exists $t \in [a - \tau, a + \tau]$ such that

$$s_k = \frac{t_k - t}{\tau} \notin \{j\kappa : j = \pm 1, \pm 2, \ldots, \pm m\}$$
$$\text{for} \qquad k = \pm 1, \pm 2, \ldots, \pm m \tag{6}$$

and

$$\left| \prod_{\substack{k=-m \\ k \neq 0}}^{m} 2s_k \right| > 1. \tag{7}$$

By Lemmas 2 and 3 we get

$$r(t; t_1, \ldots, t_n)^2 \geq \frac{2\Omega_0}{1 + \Sigma_{k=-m}^{\prime m} |b_k|^2}, \tag{8}$$

where numbers b_k satisfy

$$\sum_{k=-m}^{m}{}' \, b_k \frac{1}{s_l - k\kappa} = 1/s_l, \qquad l = 1, 2, \ldots, n. \tag{9}$$

To find b_k explicitly recall the Cauchy equality

$$\Delta(c_1, \ldots, c_n; d_1, \ldots, d_n) \equiv \det\left(\frac{1}{c_k + d_l}\right)_{k,l=1}^{n}$$
$$= \frac{\Pi_{k>l} (c_k - c_l)(d_k - d_l)}{\Pi_{k,l} (c_k + d_l)}$$

(see Natanson, 1949). This and (6) imply that (9) is a nonsingular system. By Cramer's rule we get $b_k = \mathbf{D}_k/\mathbf{D}$, where

$$\mathbf{D} = \Delta(s_1, \ldots, s_n; -m\kappa, \ldots, -\kappa, +\kappa, \ldots, m\kappa)$$

and $\mathbf{D}_k$ is derived from $\mathbf{D}$ by replacing $k\kappa$ with 0. After standard calculations we find

$$b_k = (-1)^{k+1} \frac{\binom{m}{|k|}}{\binom{m+|k|}{|k|}} \prod_{j=1}^{n} (1 - \kappa k/s_j), \qquad k = \pm 1, \pm 2, \ldots, \pm m.$$

From $\binom{m}{|k|} < \binom{m+|k|}{k}$, $|s_j| \le 2$, and (7) we see that

$$|b_k| < \frac{(4 + \kappa n)^n}{\prod_{j=1}^n |2s_j|} < (4 + \kappa n)^n, \qquad k = \pm 1, \pm 2, \ldots, \pm m.$$

These inequalities and (8) yield

$$R(t_1, t_2, \ldots, t_n) \ge r(t; t_1, \ldots, t_n) > \frac{(\Omega_0/n)^{1/2}}{(4 + \kappa n)^n},$$

which establishes the lemma for $n = 2m$.

Suppose now that $n = 2m - 1$. Then for any $t_{n+1} \in I \backslash \{t_1, t_2, \ldots, t_n\}$ we have

$$R(t_1, t_2, \ldots, t_n) \ge R(t_1, t_2, \ldots, t_n, t_{n+1}) > \frac{(\Omega_0/(n+1))^{1/2}}{(4 + \kappa(n+1))^{n+1}},$$

completing the proof. ■

Remark. The estimate in Lemma 4 is evidently not sharp for all $n > 0$. The author believes that for $n \le \kappa^{-1}$, $R(t_1, t_2, \ldots, t_n)$ is at least of order $\Omega_0^{1/2}$.

We shall see later on that for sufficiently small $\epsilon > 0$, Lemma 4 yields a sharp lower bound on comp(ϵ).

Let $l_n(\check{X})$ be the Lagrange interpolant of $\check{X} \in J_0$ with nodes $t_1, t_2, \ldots, t_n$, i.e.,

$$l_n(\check{X})(t) \equiv \sum_{j=1}^{n} \check{X}[t_1, t_2, \ldots, t_j] \prod_{k=1}^{j-1} (t - t_k), \tag{10}$$

where

$$\check{X}[t_1, t_2, \ldots, t_j] = \sum_{k=1}^{j} \check{X}(t_k) \prod_{\substack{l=1 \\ l \ne k}}^{j} (t_k - t_l)^{-1}$$

are divided differences of $\check{X}$.

Now define the algorithm Λ_n which uses information

$$N\check{X} = [\check{X}(t_1), \check{X}(t_2), \ldots, \check{X}(t_n)]^{\mathrm{T}}$$

by

$$\Lambda_n(N\check{X}) \equiv l_n(\check{X}).$$

LEMMA 5.

$$R(t_1, t_2, \ldots, t_n) \le e(\Lambda_n) < \frac{(\Omega_0/(n+\frac{1}{2}))^{1/2}}{n!\,(\kappa/2\pi)^n}.$$

Proof. We only need to prove the second inequality.

Choose any $X \in B(\Omega_0)$. Then

$$\check{X}(t) - l_n(\check{X})(t) = \check{X}[t, t_1, \ldots, t_n] \prod_{j=1}^{n} (t - t_j).$$

Since $\check{X}(t) = \int_{-\Omega_0}^{\Omega_0} X(\Omega)\exp(i\Omega t)d\Omega$ we have

$$\check{X}[t, t_1, \ldots, t_n] = \int_{-\Omega_0}^{\Omega_0} X(\Omega)(\exp(i\Omega\cdot)[t, t_1, \ldots, t_n])d\Omega.$$

Consequently,

$$\begin{aligned} e(\Lambda_n) &= \sup\{|\check{X}(t) - l_n(\check{X})(t)| : \check{X} \in J_0,\ t \in I\} \\ &= \sup\left\{\prod_{j=1}^{n} |t - t_j| \sup_{X \in B(\Omega_0)} \left| \int_{-\Omega_0}^{\Omega_0} X(\Omega)\exp(i\Omega\cdot)[t, t_1, \ldots, t_n]d\Omega \right| \right. \\ &\qquad \left. : t \in I\right\} \qquad (11) \\ &= \sup\left\{\prod_{j=1}^{n} |t - t_j| \left(\int_{-\Omega_0}^{\Omega_0} |\exp(i\Omega\cdot)[t, t_1, \ldots, t_n]|^2 d\Omega\right)^{1/2} : t \in I\right\}. \end{aligned}$$

Note now that $|\exp(i\Omega\cdot)[t, t_1, \ldots, t_n]1 \le |\Omega|^n/n!$. Thus,

$$e(\Lambda_n) \le \frac{\Omega_0^{n+1/2}}{(n+\frac{1}{2})^{1/2} n!} \sup\left\{\prod_{j=1}^{n} |t - t_j| : t \in I\right\}. \qquad (12)$$

Since the supremum in (12) is less than $(2\tau)^n$, the lemma follows easily. ∎

Remarks. (1). Note that the Stirling formula, $n! = (2\pi n)^{1/2}(n/e)^n e^{\delta/12n}$, $0 < \delta < 1$, gives

$$e(\Lambda_n) < K\frac{\Omega_0^{1/2}}{n(n\kappa_1)^n}, \qquad \text{with } \kappa_1 = \kappa/2\pi e,\ K = e^{-1/12}(2\pi)^{1/2}. \quad (13)$$

Thus for $n > \kappa_1^{-1} = 2e\Omega_0\tau$, Λ_n provides an exponentially good approximation of any $\check{X} \in J_0$, independently of the sample points' location in the interval I. In particular, even if the sample points are concentrated in an arbitrarily small subinterval $I_s \subset I$, exponentially good extrapolation of $\check{X}$ on $I\backslash I_s$ is guaranteed. To understand this, note that by the Paley–Weiner theorem J_0 consists of entire functions with the exponential type Ω_0 restricted to the real line. Hence, the above conclusion is an approximate analog of the fact that the restriction of an entire function to any interval determines the function everywhere.

We stress that good extrapolation of signals from J_0 is possible only in theory. In practice, only inaccurate samples are available. As pointed out by Kowalski and Sawoń (1983), this makes good extrapolation possible for only a bounded distance beyond the interval of observation, regardless of the number of samples used. This distance increases with the accuracy of the reading of the samples (see Kritikos, 1985).

(2) Observe now that $e(\Lambda_n)$ provides a much better estimate for a special choice of sampling points.

Namely, let Λ_n^t denote the Lagrange algorithm corresponding to the Chebyshev nodes

$$t_k^* = a + \tau\cos\frac{(2k-1)\pi}{2n}, \qquad k = 1, 2, \ldots, n.$$

Then

$$\prod_{k=1}^{n}(t - t_k^*) = 2(\tau/2)^n T_n\left(\frac{t-a}{\tau}\right),$$

where T_n is the nth Chebyshev polynomial of the first kind, $T_n(x) \equiv \cos(n \arccos x)$. Thus,

$$\sup\left\{\prod_{k=1}^{n}|t - t_k^*| : t \in I\right\} = 2(\tau/2)^n$$

and (12) gives

$$e(\Lambda_n^t) \le \frac{2(\Omega_0\tau/2)^n\Omega_0^{1/2}}{(n+\frac{1}{2})^{1/2}n!}.$$

Now applying the Stirling formula we get

$$e(\Lambda_n^t) < 2K \frac{\Omega_0^{1/2}}{n(n\kappa_2)^n}, \qquad \text{with } \kappa_2 = 4\kappa_1.$$

Thus, Λ_n^t provides an exponentially good approximation of any $\check{X} \in J_0$ for $n > \kappa_2^{-1} \approx 1.36\Omega_0\tau$.

(3) We point out that for $n < n_0 \approx (10/11)\Omega_0\tau$, $e(\Lambda_n)$ is at least of order. $0.1(\Omega_0\tau/2)^{-1/2}$, independently of the sample points chosen.

To show this, note first that

$$e(\Lambda_n) \geq (2\Omega_0\tau)^{1/2} E_{n-1}, \tag{14}$$

where $E_{n-1} = E_{n-1}(\mathrm{sa}(\Omega_0\tau\cdot))$ is the optimal error of the uniform approximation of $\mathrm{sa}(\Omega_0\tau\cdot)$ by polynomials of degree $\leq n-1$ on the interval $[-1, +1]$. Indeed, substitute in (11):

$$\Omega = \Omega_0\omega, \qquad t = u\tau + a, \qquad t_k = u_k\tau + a \qquad (k = 1, 2, \ldots, n).$$

Then ω, u, $u_k \in [-1, +1]$ and

$$|\exp(i\Omega_0\omega\cdot)[t, t_1, \ldots, t_n]| = \tau^{-n}\,|\exp(i\Omega_0\omega\tau\cdot)[u, u_1, \ldots, u_n]|.$$

Consequently, (11) takes the form

$$e(\Lambda_n) = \Omega_0^{1/2} \sup\left\{\left(\prod_{k=1}^{n} |u - u_k|\left(\int_{-1}^{1} |\exp(i\Omega_0\omega\tau\cdot)\right.\right.\right.$$
$$\left.\left.\left.[u, u_1, \ldots, u_n]|^2 d\omega\right)^{1/2} : u \in [-1, +1]\right\}.$$

By the Schwarz inequality we get

$$\begin{aligned}
&\left(2\int_{-1}^{1} |\exp(i\Omega_0\omega\tau\cdot)[u, u_1, \ldots, u_n]|^2 d\omega\right)^{1/2} \\
&\geq \left|\int_{-1}^{1} \exp(i\Omega_0\omega\tau\cdot)[u, u_1, \ldots, u_n] d\omega\right| \\
&= \left|\left(\int_{-1}^{1} \exp(i\Omega_0\omega\tau\cdot) d\omega\right)[u, u_1, \ldots, u_n]\right| \\
&= 2\,|\mathrm{sa}(\Omega_0\tau\cdot)[u, u_1, \ldots, u_n]|.
\end{aligned}$$

Hence,

$$(2\Omega_0)^{-1/2}e(\Lambda_n) \geq \sup\left\{\left|\prod_{k=1}^{n}(u-u_k)\mathrm{sa}(\Omega_0\tau\cdot)[u,u_1,\ldots,u_n]\right| : u \in [-1,+1]\right\}$$
$$= \sup\{|\mathrm{sa}(\Omega_0\tau u) - \mathrm{ln}(\mathrm{sa}(\Omega_0\tau\cdot))(u)| : u \in [-1,+1]\}$$
$$\geq E_{n-1},$$

which proves (14).

Suppose now that $h = \lfloor(n-1)/2\rfloor$ satisfies

$$\frac{1.1\pi}{\Omega_0\tau\sin(\pi/2(h+1))} \equiv s \leq 1. \tag{15}$$

Observe that for sufficiently large $\Omega_0\tau$ this means that $n \leq 2h+1$ and $h \leq m \approx (5/11)\Omega_0\tau$. Define

$$P_{2m}(x) \equiv \frac{(-1)^m}{2(m+1)}\frac{U_{2m+1}(s^{-1}x)}{s^{-1}x} \quad \text{and} \quad \delta(x) \equiv \mathrm{sa}(\Omega_0\tau x) - p_{2m}(x),$$

where U_{2m+1} is the $(2m+1)$st Chebyshev polynomial of the second kind, i.e., $U_{2m+1}(z) = \sin[2(m+1)\arccos z])(1-z^2)^{-1/2}$. Since U_{2m+1} is odd, p_{2m} is a polynomial of degree $2m$.

Note that $\delta(0) = 0$ and for $q_j = s\cos((2j-1)\pi/4(m+1))$ $(j = 1, 2, \ldots, 2m+2)$ we have $-1 < q_{2(m+1)} < \ldots < q_{m+2} < 0 < q_{m+1} < \ldots < q_1 < 1$ and

$$p_{2m}(q_j) = \frac{(-1)^{m+j+1}}{(m+1)\sin((2j-1)\pi/2(m+1))}.$$

Moreover,

$$|\mathrm{sa}(\Omega_0\tau q_j)| \leq 1/|\Omega_0\tau q_j| \leq \frac{10}{11}\frac{1}{2(m+1)\,|\cos((2j-1)\pi/4(m+1))|}$$
$$= \frac{10}{11}\frac{|\sin((2j-1)\pi/4(m+1))|}{(m+1)\,|\sin((2j-1)\pi/2(m+1))|} \leq \frac{10}{11}|p_{2m}(q_j)|.$$

Consequently,

$$\mathrm{sgn}\,\delta(q_j) = \mathrm{sgn}\,p_{2m}(q_j) = \begin{cases}(-1)^{m+j+1} & \text{if } 1 \leq j \leq m+1,\\ (-1)^{m+j} & \text{if } m+2 \leq j \leq 2m+2\end{cases}$$

and

$$|\delta(q_j)| \geq |p_{2m}(q_j)| - |\mathrm{sa}(\Omega_0 \tau q_j)| \geq \frac{1}{11}|p_{2m}(q_j)| \geq \frac{1}{11(m+1)}. \quad (16)$$

The Remez theorem now implies $E_{2m} \geq B \equiv \max_j(|\delta(\overline{q}_j)| + |\delta(\overline{q}_{j+1})|)/2$, where

$$\overline{q}_j = \begin{cases} q_j & \text{if } 1 \leq j \leq m+1, \\ 0 & \text{if } j = m+2, \\ q_{j-1} & \text{if } m+3 \leq j \leq 2(m+1) \end{cases}$$

(see Meinardus, 1964, or Remez, 1934). By (14), (15), (16), and the obvious inequality $E_{n-1} \geq E_{2m}$ we finally get

$$e(\Lambda_n) \geq (2\Omega_0\tau)^{1/2}B \geq \frac{1}{11}\frac{(2\Omega_0\tau)^{1/2}}{2(m+1)} \geq \frac{1}{11}(2\Omega_0\tau)^{1/2}\frac{\arcsin 1.1\pi/(\Omega_0\tau)}{\pi}$$
$$\approx 0.1(\Omega_0\tau/2)^{-1/2},$$

as claimed.

3. Proof of Theorem 1

As follows from Lemmas 4 and 5 and (13), for any $\kappa_3 > \kappa$ and sufficiently large n we have

$$K\frac{\Omega_0^{1/2}}{n(n\kappa_1)^n} > e(\varphi^*) > \frac{(\Omega_0/n)^{1/2}}{(n\kappa_3)^n}, \quad (17)$$

where φ^* is an optimal error algorithm that uses n samples. Thus, if $\epsilon > 0$ is sufficiently small, $m(\epsilon)$ must satisfy

$$K_1\frac{\Omega_0^{1/2}}{m(\epsilon)(m(\epsilon)\kappa_1)^{m(\epsilon)}} > \epsilon > \frac{(\Omega_0/m(\epsilon))^{1/2}}{(m(\epsilon)\kappa_3)^{m(\epsilon)}},$$

for some $K_1 \geq K$. This yields

$$m(\epsilon)\,(\log m(\epsilon) + \log \kappa_1 + o(1)) < \log 1/\epsilon$$
$$< m(\epsilon)\,(\log m(\epsilon) + \log \kappa_3 + o(1)).$$

Substituting $m(\epsilon) = h(\epsilon)\log 1/\epsilon/\log\log 1/\epsilon$ we get

$$h(\epsilon)\left(1 - \frac{\log\log\log 1/\epsilon}{\log\log 1/\epsilon} + 0\left(\frac{\log \kappa_1}{\log\log 1/\epsilon}\right)\right) < 1$$

$$< h(\epsilon)\left(1 - \frac{\log\log\log 1/\epsilon}{\log\log 1/\epsilon} + 0\left(\frac{\log \kappa_3}{\log\log 1/\epsilon}\right)\right).$$

Consequently, $h(\epsilon) = 1 + o(1)$ and (i) of Theorem 1 follows easily.

Let $m_1(\epsilon)$ denote the minimal number of nodes from I such that the corresponding Lagrange algorithm yields an ϵ-approximation, regardless of their location in I, i.e.,

$$m_1(\epsilon) = \min\{m : e(\Lambda_m) \leq \epsilon, \forall\, t_1, t_2, \ldots, t_m \in I\}.$$

By Lemmas 4 and 5, if n is sufficiently large, (17) holds for φ^* replaced by Λ_n. Proceeding as in the proof of (i) we get

$$m_1(\epsilon) = m(\epsilon)(1 + o(1)) = \Theta\left(\frac{\log 1/\epsilon}{\log\log 1/\epsilon}\right). \tag{18}$$

Recall now that (10) can be rewritten in the form

$$\Lambda_n(N\check{X})(t) = \begin{cases} \mathcal{P}(t) \displaystyle\sum_{k=1}^{n} \frac{\check{X}(t_k)}{\mathcal{P}'(t_k)(t - t_k)} & \text{if } t \neq t_k, \\ \check{X}(t_k) & \text{if } t = t_k, \end{cases}$$

where $\mathcal{P}(t) = \Pi_{k=1}^{n} (t - t_k)$. Note that numbers $\mathcal{P}'(t_k)$ can be precomputed. Then, to get $\Lambda_n(N\check{X})(t)$ we need n measurements of $\check{X}$ and at most $3n$ arithmetic operations.

Denote by $\text{comp}(\Lambda_{m_1(\epsilon)})$ the total cost of producing an ϵ-approximation by $\Lambda_{m_1(\epsilon)}$. Then

$$cm(\epsilon) \leq \text{comp}(\epsilon) \leq \text{comp}(\Lambda_{m_1(\epsilon)}) \leq cm_1(\epsilon) + 3m_1(\epsilon).$$

Thus, (i) and (18) imply that $\text{comp}(\epsilon)$ and $\text{comp}(\Lambda_{m_1(\epsilon)})$ are of order

$$\frac{\log 1/\epsilon}{\log\log 1/\epsilon} \quad \text{as } \epsilon \to 0^+.$$

This proves (ii) and (iii), completing the proof of Theorem 1. ■

Acknowledgments

I am indebted to W. Ogryczak for calling my attention to Golshtein (1971) and Hestens (1966). I also wish to thank A. G. Werschulz and H. Woźniakowski for reading the manuscript and for valuable remarks.

REFERENCES

BUTZER, P. L. (1983), A survey of Whittaker–Shannon theorem and some of its extensions, *J. Math. Res. Exposition* **3**, 185–212.

GOLSHTEIN, E. G. (1971), "Duality Theory in Mathematical Programming and Its Applications," Nauka, Moscow. [In Russian]

HESTENS, M. (1966), "Calculus of Variations and the Optimal Control Theory," Wiley, New York.

KOWALSKI, M. A., AND SAWOŃ, Z. (1983), The moment problem in the space $C_0(S)$, *Mh. Math.* **97**, 47–53.

KREIN, M. G., AND NUDELMAN, A. A. (1973), "The Markov Moment Problem and Extremal Problems," Nauka, Moscow. [In Russian]

KRITIKOS, H. (1985), Measurement accuracies in bandlimited extrapolation, to appear.

LANDAU, H. J. (1985), "Extrapolating a Bandlimited Function from Its Samples Taken in a Finite Interval," AT&T Bell Laboratories.

MEINARDUS, G. (1964), "Approximation von Functionen und ihre numerische Behandlung," Springer-Verlag, Berlin.

MICCHELLI, C. A., AND RIVLIN, T. J. (1977), A survey of optimal recovery, *in* "Optimal Estimation in Approximation Theory" (C. A. Micchelli and T. J. Rivlin, Eds.), Plenum, New York.

NATANSON, J. P. (1949), "Constructive Theory of Functions," Leningrad. [In Russian]

REMEZ, E. J. (1934), Sur un procédé convergent d'approximations successives pour déterminer les polynômes d'approximation, *C. R. Acad. Sci. Paris* **198**, 2063–2065.

SAKS, S., AND ZYGMUND, A. (1971), "Analytic Functions," 3rd ed., Elsevier/PWN, Amsterdam/London/New York.

TRAUB, J. F., AND WOŹNIAKOWSKI, H. (1980), "A General Theory of Optimal Algorithms," Academic Press, New York.